

97 FEB 2003

PCT/JP03/09483

日 本 国 特 許 庁

JAPAN PATENT OFFICE

25.07.03

REC'D 12 SEP 2003

WIPO PAT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日  
Date of Application: 2002年 8月 7日

出 願 番 号  
Application Number: 特願2002-230312,  
[ST. 10/C]: [JP2002-230312]

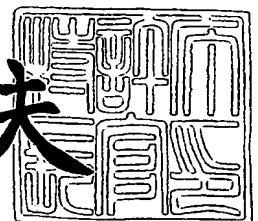
出 願 人  
Applicant(s): 松下電器産業株式会社

**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

2003年 8月28日

特許庁長官  
Commissioner,  
Japan Patent Office

今井康夫



Best Available Copy

【書類名】 特許願

【整理番号】 2903240058

【提出日】 平成14年 8月 7日

【あて先】 特許庁長官殿

【国際特許分類】 G06K 9/46

【発明者】

    【住所又は居所】 神奈川県横浜市港北区綱島東四丁目 3 番 1 号 松下通信工業株式会社内

    【氏名】 藤田 幸宏

【発明者】

    【住所又は居所】 神奈川県横浜市港北区綱島東四丁目 3 番 1 号 松下通信工業株式会社内

    【氏名】 沢村 恒治

【特許出願人】

    【識別番号】 000005821

    【氏名又は名称】 松下電器産業株式会社

【代理人】

    【識別番号】 100105647

    【弁理士】

    【氏名又は名称】 小栗 昌平

    【電話番号】 03-5561-3990

【選任した代理人】

    【識別番号】 100105474

    【弁理士】

    【氏名又は名称】 本多 弘徳

    【電話番号】 03-5561-3990

## 【選任した代理人】

【識別番号】 100108589

【弁理士】

【氏名又は名称】 市川 利光

【電話番号】 03-5561-3990

## 【選任した代理人】

【識別番号】 100115107

【弁理士】

【氏名又は名称】 高松 猛

【電話番号】 03-5561-3990

## 【選任した代理人】

【識別番号】 100090343

【弁理士】

【氏名又は名称】 栗宇 百合子

【電話番号】 03-5561-3990

## 【手数料の表示】

【予納台帳番号】 092740

【納付金額】 21,000円

## 【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0002926

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 文字認識処理装置及び文字認識処理方法

【特許請求の範囲】

【請求項 1】 文字認識用に撮影された認識対象となる文字画像の画像データを取り込む画像取り込み手段と、

前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するカーソル情報出力手段と、

前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合して文字の配置を解析するレイアウト解析手段と、

前記レイアウト解析手段の解析結果を基に 1 文字分の文字画像を切り出す文字切り出し手段と、

前記切り出された 1 文字分の文字画像を文字として認識し文字情報に変換する文字認識手段と、

を備えたことを特徴とする文字認識処理装置。

【請求項 2】 前記画像取り込み手段は、認識対象となる文字画像が連続撮影により得られた複数の文字画像からなる場合に、この文字画像全体から所定領域ずつ文字画像の画像データを分割して取り込み、

前記レイアウト解析手段は、前記取り込んだ複数の文字画像の画像データごとに、前記カーソル位置情報を照合して文字の配置を解析することを特徴とする請求項 1 に記載の文字認識処理装置。

【請求項 3】 請求項 1 または 2 に記載の文字認識処理装置を備えた携帯端末装置。

【請求項 4】 認識対象となる文字画像を文字認識用文字枠とともに表示して撮影するステップと、

前記撮影した文字画像の画像データを取り込むステップと、

前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するステップと、

前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合して文字の配置を解析するステップと、

前記文字配置の解析結果を基に 1 文字分の文字画像を切り出すステップと、  
前記切り出された 1 文字分の文字画像を文字として認識し文字情報に変換する  
ステップと、

を有することを特徴とする文字認識処理方法。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、携帯電話機や携帯情報端末等の携帯端末装置に用いて好適な文字認識処理装置及び文字認識処理方法に関する。

【0 0 0 2】

【従来の技術】

従来の文字認識方法及び装置として、例えば特開 2 0 0 1 - 6 0 2 5 0 号公報  
や特開 2 0 0 0 - 7 6 3 7 8 号公報に開示されているものが知られている。

【0 0 0 3】

特開 2 0 0 1 - 6 0 2 5 0 号公報に開示された文字認識方法及び装置は、画像  
入力された文字から特徴量を抽出し、抽出した特徴量に基づいて文字の認識を行  
うものである。この従来例では、画像入力された文字を所定間隔で走査する毎に  
走査線と文字との交差回数を取得するとともに、取得した交差回数の隣接する走  
査線毎の差分の総和を特徴量の 1 つとして文字の認識を行うものである。認識対  
象となる文字に対する走査は、少なくとも 2 つの異なる方向に所定間隔に行った  
り、直交する 2 方向に所定間隔で行うようにしている。

【0 0 0 4】

他方、特開 2 0 0 0 - 7 6 3 7 8 号公報に開示されている文字認識方法は、入  
力された画像データの輪郭線データを作成し、作成した輪郭線データを正規化し  
て得たデータを、予め登録している標準文字の輪郭線データを抽出して作成した  
辞書データに照合して文字認識を行うものである。この際、入力された画像デー  
タの書体や歪などによっては、辞書データとの照合前に画像データの輪郭線デー  
タを辞書データで認識できる書体に変形したり、あるいは歪が是正されるように  
変形するようにしている。

## 【0005】

## 【発明が解決しようとする課題】

しかしながら、従来の文字認識方法及び装置においては、取り込んだ文字画像データから文字認識処理を行う際に、認識率や認識速度の向上を図るためには、レイアウト解析や行、文字の切り出し作業においてそれぞれ走査線や輪郭線データ等を用いて特別な処理を行う必要があった。

## 【0006】

また、一般に文字認識を行う文字列全体を一度に撮影する場合、1文字に対する解像度を確保するには撮影装置（例えばCCDカメラなど）の画素数を大きくしなければならない上、必要以上に大規模なデータを扱うために、処理に必要なホストCPUやDSPのMIPS (Million Instructions Per Second) 値やワークRAMを大きくしなければならないという問題点もある。

## 【0007】

本発明は、上記事情に鑑みてなされたもので、その目的は、認識率及び認識速度の向上を図りつつ、撮影装置やシステムを構成するデバイスに必要とされる仕様の程度を低減でき、低消費電力化及びコストダウンを図ることが可能な文字認識処理装置及び文字認識処理方法を提供することにある。

## 【0008】

## 【課題を解決するための手段】

本発明に係る文字認識処理装置は、文字認識用に撮影された認識対象となる文字画像の画像データを取り込む画像取り込み手段と、前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するカーソル情報出力手段と、前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合して文字の配置を解析するレイアウト解析手段と、前記レイアウト解析手段の解析結果を基に1文字分の文字画像を切り出す文字切り出し手段と、前記切り出された1文字分の文字画像を文字として認識し文字情報に変換する文字認識手段と、を備えたことを特徴とする。

## 【0009】

上記構成では、文字認識対象物の文字認識を行う際に、認識対象となる文字画

像を文字認識用文字枠とともに液晶表示素子等による表示部に表示して、認識対象文字列に文字認識用文字枠のカーソルを合わせた状態で文字画像をCCDカメラ等の撮像部により撮影する。そして、画像取り込み手段で文字画像の画像データを取り込み、レイアウト解析手段により取り込んだ文字画像の画像データに対してカーソル位置情報を照合して行や文字等の配置を解析し、文字切り出し手段及び文字認識手段によって文字画像に対する文字の切り出し及び文字認識を行う。このため、レイアウト解析手段においてカーソル位置情報を利用して認識対象文字列の行や文字の配置情報を正確に解析することが可能となる。これにより、認識率及び認識速度の向上が図れる。また、カーソル位置情報を利用することで文字認識用の画像データの画素数を多くしなくても精度の高い文字認識が可能であるため、撮影部を多画素化するなどの高性能の仕様は必要なく、装置に要求されるスペックの軽減が図れる。さらに、装置内でのデータ処理量を低減することができ、これによって、CPUやDSP、メモリ等のシステムを構成するデバイスに必要とされるスペックを軽減できるため、装置各部の仕様の程度を低減でき、低消費電力化及びコストダウンを図れる。

#### 【0010】

また、前記画像取り込み手段は、認識対象となる文字画像が連続撮影により得られた複数の文字画像からなる場合に、この文字画像全体から所定領域ずつ文字画像の画像データを分割して取り込み、前記レイアウト解析手段は、前記取り込んだ複数の文字画像の画像データごとに、前記カーソル位置情報を照合して文字の配置を解析することを特徴とする。

#### 【0011】

上記構成では、文字認識対象物の認識対象文字列が大きい場合など、1度の撮影では文字画像全体を撮影できない場合に、撮影部を移動させながら連続撮影を行って文字画像を撮影する。そして、得られた文字画像の画像データについて、画像取り込み手段で文字画像全体から所定領域ずつ文字画像の画像データを分割して取り込み、レイアウト解析手段により取り込んだ複数の文字画像の画像データごとにカーソル位置情報を照合して行や文字等の配置を解析する。このように文字画像をいくつかのコマの静止画像に分けてカーソル位置情報とともに取り込

むことにより、認識対象文字列が長い場合などでもレイアウト解析手段においてカーソル位置情報を利用して認識対象文字列の配置情報を正確に解析することが可能となる。

#### 【0012】

また本発明は、上記いずれかに記載の文字認識処理装置を備えた携帯端末装置を提供する。

#### 【0013】

上記構成によれば、文字認識処理装置を搭載した携帯端末装置において、文字認識の認識率及び認識速度の向上を図りつつ、撮影部やCPU、DSP、メモリ等のシステムを構成するデバイスに必要とされるスペックを軽減できるため、装置各部の仕様の程度を低減でき、低消費電力化及びコストダウンを図れる。

#### 【0014】

本発明に係る文字認識処理方法は、認識対象となる文字画像を文字認識用文字枠とともに表示して撮影するステップと、前記撮影した文字画像の画像データを取り込むステップと、前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するステップと、前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合して文字の配置を解析するステップと、前記文字配置の解析結果を基に1文字分の文字画像を切り出すステップと、前記切り出された1文字分の文字画像を文字として認識し文字情報に変換するステップと、を有することを特徴とする。

#### 【0015】

この方法によれば、カーソル位置情報を利用して認識対象文字列の行や文字の配置情報を正確に解析することが可能となり、認識率及び認識速度の向上が図れる。また、カーソル位置情報を利用することにより、装置各部に要求されるスペックの軽減が図れるとともに、装置内でのデータ処理量を低減できるため、装置の仕様の程度を低減でき、低消費電力化及びコストダウンを図ることが可能となる。

#### 【0016】

#### 【発明の実施の形態】



以下、図面を参照して本発明の実施形態を説明する。

(第1実施形態)

図1は、本発明の第1実施形態に係る文字認識処理装置を備えた携帯端末装置の構成を示すブロック図である。本実施形態では、文字認識処理装置2を携帯端末装置1に搭載した構成例を説明する。

【0017】

本実施形態の文字認識処理装置2は、画像取り込み部3と、カーソル制御部4と、レイアウト解析部5と、文字切り出し部6と、文字認識部7と、フォーマット出力部8とを備えて構成される。

【0018】

携帯端末装置1は、文字認識処理装置2とともに、例えばCCDカメラ等からなる撮影部9と、表示制御部10と、アプリケーション部11とを有している。撮影部9は、文字認識対象物を撮影して撮影データとして取り込み、文字認識処理装置2の画像取り込み部3に入力する。画像取り込み部3は、画像取り込み手段に相当するもので、入力された撮影データを所定のフォーマットの画像データに変換して表示制御部10に入力する。カーソル制御部4は、カーソル情報出力手段に相当するもので、文字認識用文字枠として機能するカーソルに関するカーソル位置情報を表示制御部10に入力する。表示制御部10は、入力された画像データとカーソル位置情報とを表示データに変換し、さらに液晶表示素子等からなるディスプレイ12に適したフォーマットに加工して出力し、ディスプレイ12において情報表示を行う。

【0019】

レイアウト解析部5は、レイアウト解析手段に相当するもので、文字認識時に画像取り込み部3から所定フォーマットの画像データを受け取るとともに、カーソル制御部4からカーソル位置情報を受け取り、カーソル位置情報及び画像データを照合してレイアウト解析を行う。文字切り出し部6は、文字切り出し手段に相当するもので、レイアウト解析結果を元に画像データを文字単位で切り出して抽出する。文字認識部7は、文字認識手段に相当するもので、切り出された文字単位の画像データの認識処理を行って文字データに変換する。フォーマット出力部

8は、文字認識部7で認識された文字データを適切なフォーマットに変換する。アプリケーション部11は、フォーマット出力部8でフォーマット変換された文字データをアプリケーション用のデータとして利用し、必要に応じて表示制御部10に入力してディスプレイ12に表示する。

#### 【0020】

次に、図2を参照して第1実施形態の文字認識処理装置の動作について説明する。図2は第1実施形態に係る文字認識処理時の動作説明図であり、認識対象文字列を撮影している状態を示している。

#### 【0021】

携帯端末装置1の撮影部9によって文字認識対象物16の中の認識対象文字列17を撮影し、文字認識を行う。撮影している認識対象文字列17をディスプレイ12に表示する際に、認識対象文字列17と併せて文字認識用文字枠であるカーソル18が表示される。使用者は撮影位置を調整しながら認識対象文字列17をカーソル18に合わせ、この状態で操作ボタン19を操作して静止画像の撮影を実行する。

#### 【0022】

撮影された認識対象文字列17の静止画像は、画像取り込み部3で定められたフォーマットの画像データに変換されて取り込まれ、レイアウト解析部5に渡される。また、レイアウト解析部5ではカーソル制御部4からカーソル位置情報を受け取り、カーソル位置情報と文字画像の画像データとを照合する。このように、レイアウト解析部5では、カーソル位置情報を利用して認識対象文字列17の行及び文字の配置情報を正確に解析することができる。そして、文字切り出し部6で画像データの文字単位での切り出しが行われ、文字認識部7において抽出された文字単位の画像データが文字データに変換され、文字認識が行われる。

#### 【0023】

このように、第1実施形態の文字認識処理装置では、文字認識対象物16の認識対象文字列17を文字認識する際に、ディスプレイ12に認識対象文字列17の画像と併せて文字認識用文字枠のカーソル18を表示させておき、使用者が認識対象文字列17にカーソル18を合わせた状態で静止画像の撮影を行い、撮影

された文字画像の画像データをカーソル位置情報とともに取り込むようにする。これにより、レイアウト解析部 5 においてカーソル位置情報を利用して行や文字の配置情報を容易かつ正確に解析することができ、文字認識における認識率及び認識速度の向上を図ることができる。

#### 【0024】

##### (第 2 実施形態)

次に、図 3 及び図 4 を用いて本発明の第 2 実施形態に係る文字認識処理装置の動作について説明する。図 3 及び図 4 は第 2 実施形態に係る文字認識処理時の動作説明図であり、図 3 は認識対象文字列を撮影している状態を示し、図 4 は複数に分割した認識対象文字列に対する文字認識処理の様子を示している。なお、文字認識処理装置の構成は第 1 実施形態と同様であるので、図 1 を援用するものとする。

#### 【0025】

第 2 実施形態では、第 1 実施形態と同様、携帯端末装置 1 の撮影部 9 によって文字認識対象物 22 の中の認識対象文字列 23 を撮影して文字認識を行う。この場合、認識対象文字列 23 の範囲が広いために 1 度に取り込めない場合は以下のように操作する。

#### 【0026】

撮影している認識対象文字列 23 をディスプレイ 12 に表示する際に、認識対象文字列 23 と併せて文字認識用文字枠であるカーソル 18 が表示される。使用者は撮影位置を調整しながら認識対象文字列 23 の開始端位置の部分（図 3 の例では左端の部分）をカーソル 18 に合わせ、この状態で操作ボタン 19 を操作して連続静止画像撮影を開始する。撮影開始後、使用者は携帯端末装置 1 の移動方向を固定しながら、文字の筆記方向に向かうように撮影開始時の開始端位置と逆方向（図 3 の例では右方向）に認識対象文字列 23 に対して平行にスライドさせ、認識対象文字列 23 の終了端位置（図 3 の例では右端位置）までカーソル 18 に収まるようにディスプレイ 12 に表示させたところで連続静止画像撮影を終了する。このとき、使用者は例えば連続静止画像撮影の最中は操作ボタン 19 を押下したままにする。

## 【0027】

撮影された認識対象文字列 23 の連続静止画像は、画像取り込み部 3 で図 4 のように複数の画像データに分割されて取り込まれる。この場合、文字画像全体から 1 行分の文字画像についていくつかのコマ (1) ~ (4) の静止画像に分割されて抽出された後、定められたフォーマットの画像データに変換されてレイアウト解析部 5 に渡される。

## 【0028】

また、レイアウト解析部 5 ではカーソル制御部 4 からカーソル位置情報を受け取り、カーソル位置情報と文字画像の画像データとを照合する。但し、レイアウト解析に利用する各コマの文字画像に関してそれぞれの文字がカーソルの文字枠に収まっていない可能性があるため、カーソル位置情報としては「行」に対する配置情報の解析には問題なく利用することができるが、文字ごとの配置情報の解析には利用できない場合がある。この場合、カーソル位置情報を連続静止画像の移動方向にシフトさせることによって各文字の配置情報を解析することも可能である。このように、レイアウト解析部 5 においては、認識対象文字列 23 が長い場合でもカーソル位置情報を利用して認識対象文字列 23 の配置情報を正確に解析することができる。本実施形態では、行の配置情報について正確な解析が可能である。

## 【0029】

そして、文字切り出し部 6 で文字単位の切り出しが行われ、文字認識部 7 において抽出された文字単位の画像データが文字データに変換され、文字認識が行われる。この場合、文字認識結果として、コマ (1) では「A B C D E F G H I J K L」が得られ、コマ (2) では「F G H I J K L M N O P」が得られ、コマ (3) では「J K L M N O P Q R S T U」が得られ、コマ (4) では「O P Q R S T U V W」が得られる。文字認識部 7 では、これらのコマ毎の文字認識結果を合成することにより、最終文字認識結果として「A B C D E F G H I J K L M N O P Q R S T U V W」が得られる。

## 【0030】

なお、第 2 実施形態では、簡単のため 1 行分の文字画像をカーソルと照合して

一連の文字認識の処理をするかのごとく説明したが、図3に示すようにカーソル18が2行分存在する場合、一連の文字認識の処理についても2行分同時に行うことができる。

#### 【0031】

このように、第2実施形態の文字認識処理装置では、文字認識対象物22の認識対象文字列23を文字認識する際に、ディスプレイ12に認識対象文字列17の画像と併せて文字認識用文字枠のカーソル18を表示させておき、使用者が携帯端末装置1を移動させながら連続静止画像撮影を行う。そして、撮影した文字画像をいくつかの静止画像に分けて、各コマの画像データをカーソル位置情報とともに取り込むようにする。これにより、レイアウト解析部5においてカーソル位置情報を利用して行の配置情報を容易かつ正確に解析することができ、文字認識における認識率及び認識速度の向上を図ることができる。

#### 【0032】

上述したように本実施形態では、カーソル位置情報を利用することで文字認識用の画像データの画素数を多くしなくても精度の高い文字認識が可能であるため、撮影部9を多画素化するなどの高性能の仕様は必要なく、装置に要求されるスペックの軽減が図れる。さらに、装置内でのデータ処理量を低減することができ、これによって、CPUやDSP、メモリ等のシステムを構成するデバイスに必要とされるスペックを軽減できる。したがって、装置各部の仕様の程度を低減でき、低消費電力化及びコストダウンを図れる。

#### 【0033】

なお、本発明は上述した実施形態に何ら限定されるものではなく、その要旨を逸脱しない範囲において種々の態様で実施し得るものである。

#### 【0034】

##### 【発明の効果】

以上説明したように本発明によれば、認識率及び認識速度の向上を図りつつ、撮影装置やシステムを構成するデバイスに必要とされる仕様の程度を低減でき、低消費電力化及びコストダウンを図ることが可能な文字認識処理装置及び文字認識処理方法を提供できる。

**【図面の簡単な説明】****【図 1】**

本発明の第 1 実施形態に係る文字認識処理装置を備えた携帯端末装置の構成を示すブロック図

**【図 2】**

本発明の第 1 実施形態に係る文字認識処理装置における文字認識処理時の動作説明図であり、認識対象文字列を撮影している状態を示す図

**【図 3】**

本発明の第 2 実施形態に係る文字認識処理装置における文字認識処理時の動作説明図であり、認識対象文字列を撮影している状態を示す図

**【図 4】**

本発明の第 2 実施形態に係る文字認識処理装置における文字認識処理時の動作説明図であり、複数に分割した認識対象文字列に対する文字認識処理の様子を示す図

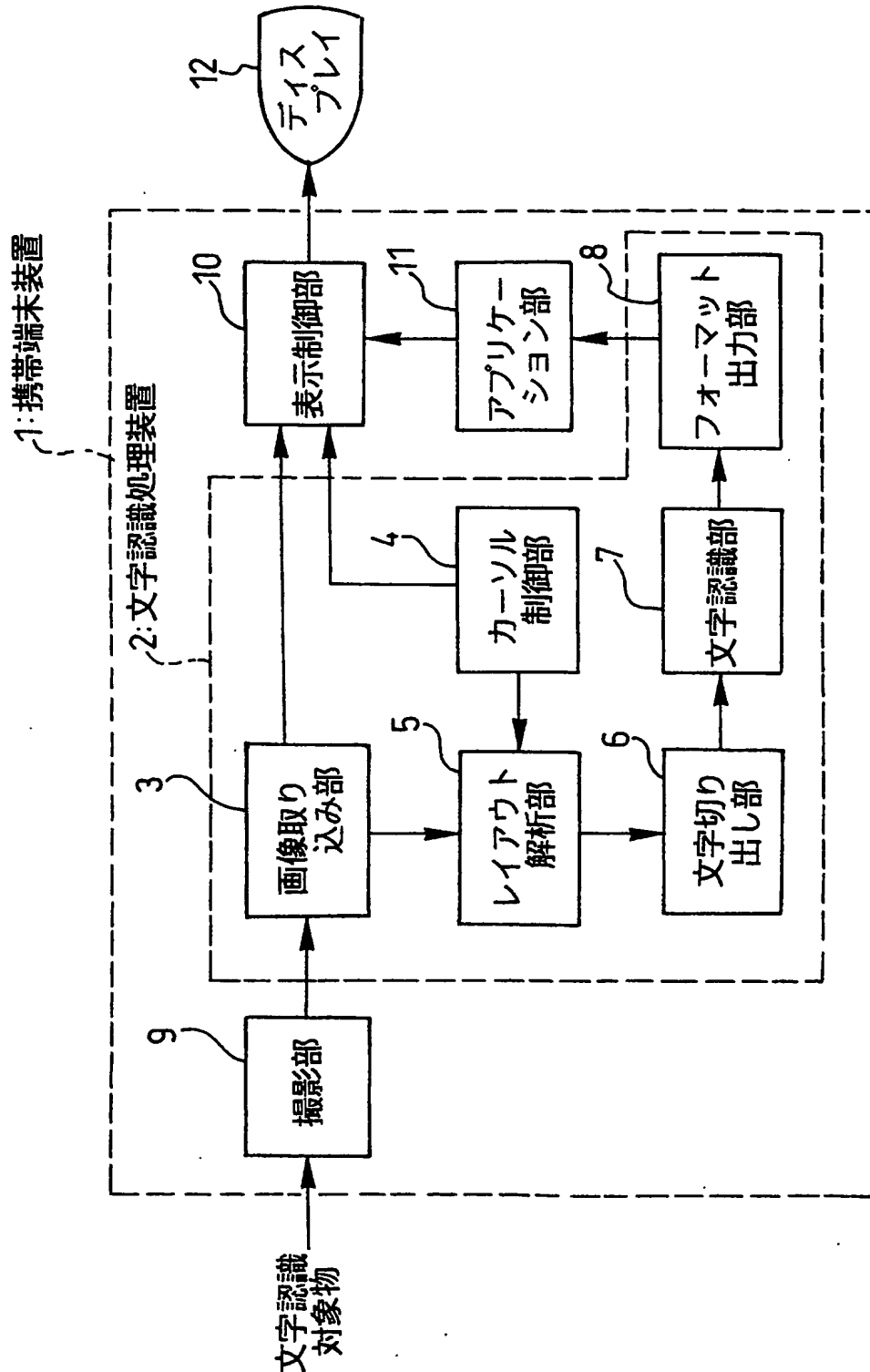
**【符号の説明】**

- 1 携帯端末装置
- 2 文字認識処理装置
- 3 画像取り込み部
- 4 カーソル制御部
- 5 レイアウト解析部
- 6 文字切り出し部
- 7 文字認識部
- 8 フォーマット出力部
- 9 撮影部
- 10 表示制御部
- 11 アプリケーション部
- 12 ディスプレイ

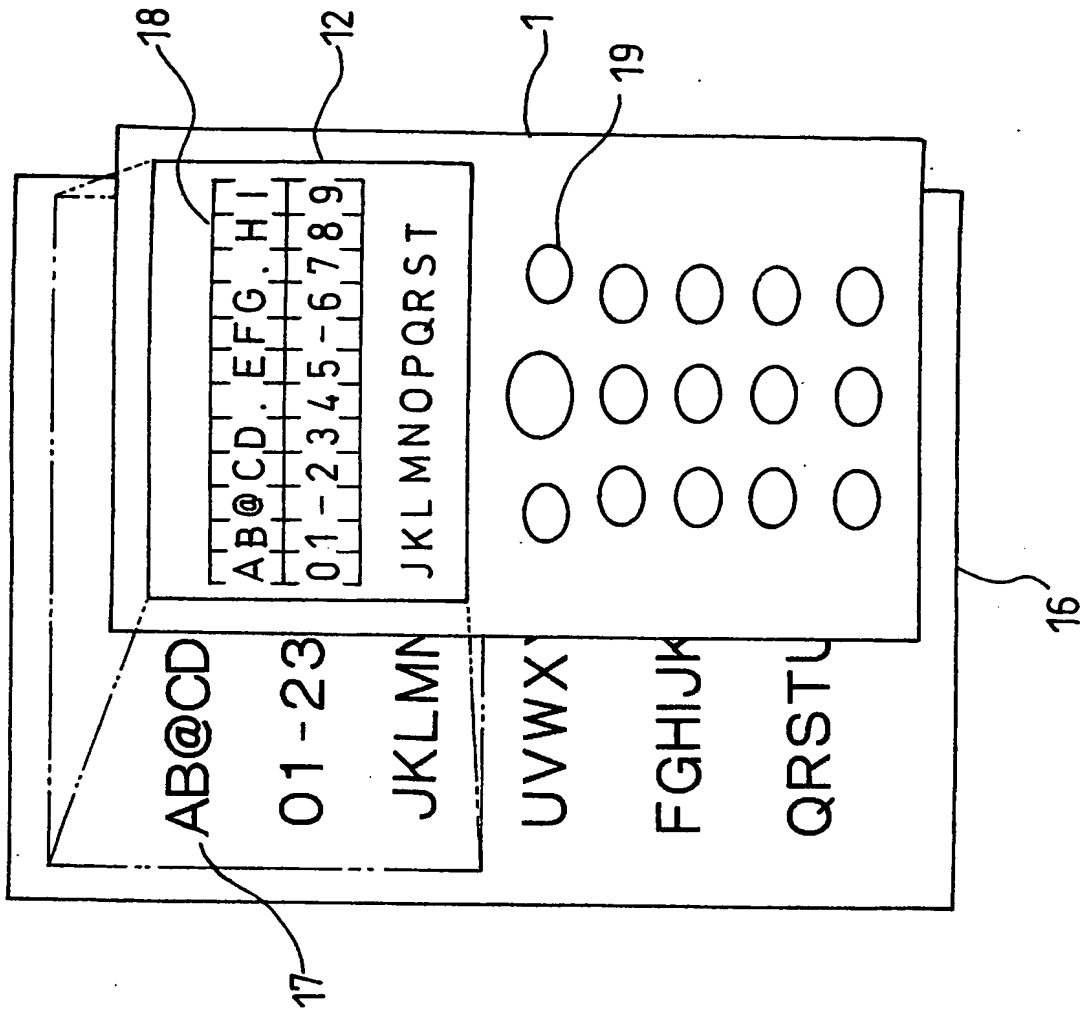
【書類名】

図面

【図 1】

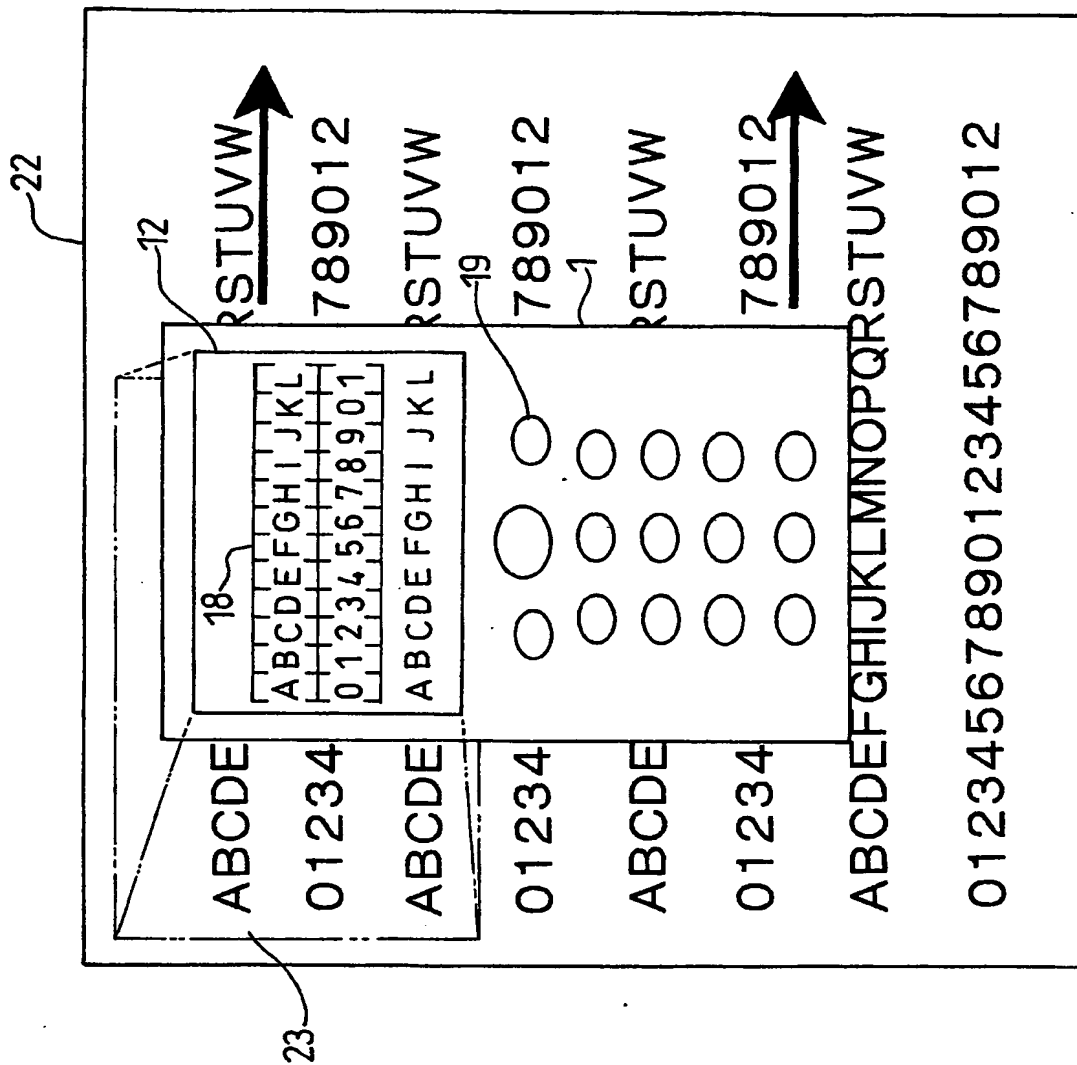


【図 2】

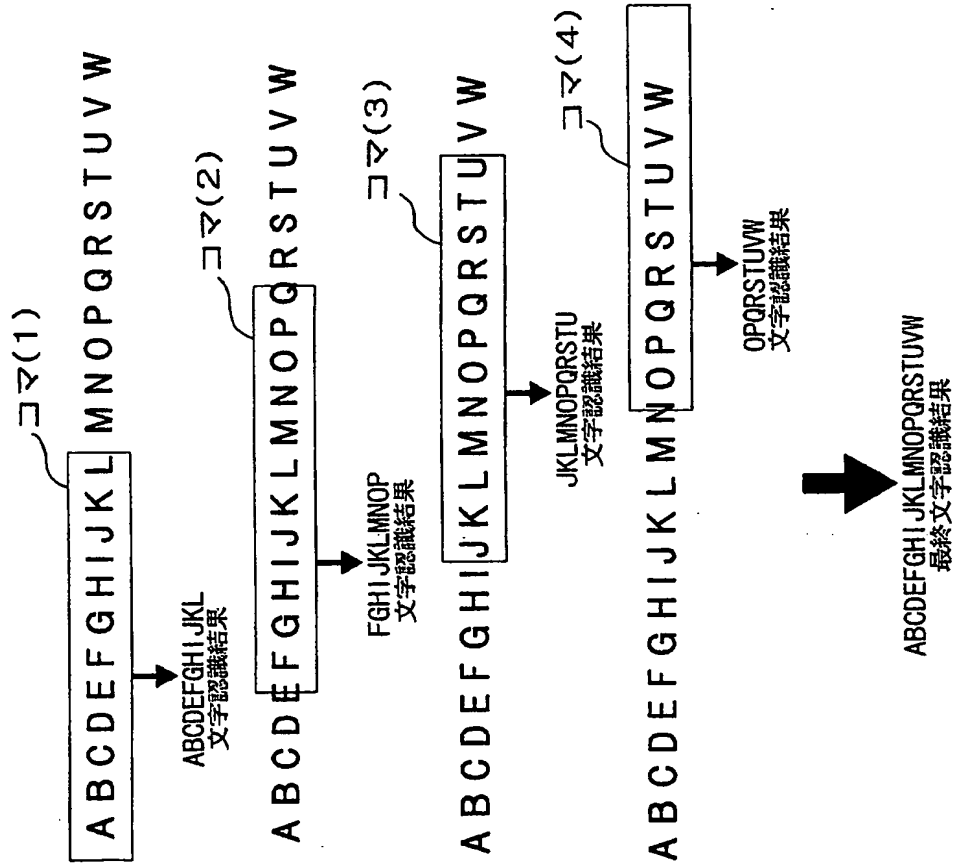




【図 3】



【図 4】



【書類名】 要約書

【要約】

【課題】 認識率及び認識速度の向上を図りつつ、撮影装置やシステムを構成するデバイスに必要とされる仕様の程度を低減可能とし、低消費電力化及びコストダウンを図る。

【解決手段】 文字認識対象物の文字画像を文字認識用文字枠のカーソルとともにディスプレイ 1 2 に表示し、認識対象文字列にカーソルを合わせた状態で撮影部 9 により撮影する。画像取り込み部 3 において、撮影された文字画像の画像データを取り込み、レイアウト解析部 5 において、取り込んだ文字画像の画像データとこの文字画像に対応するカーソル制御部 4 からのカーソル位置情報とを受け取り、画像データとカーソル位置情報とを照合して認識対象文字列の行や文字の配置を解析する。そして、文字切り出し部 6 で解析結果を基に文字の切り出しを行い、文字認識部 7 によって文字画像に対する文字認識を行って画像データを文字データに変換する。

【選択図】 図 1

特願2002-230312

出願人履歴情報

識別番号

[000005821]

1. 変更年月日

1990年 8月28日

[変更理由]

新規登録

住 所

大阪府門真市大字門真1006番地

氏 名

松下電器産業株式会社

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☒ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☒ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☒ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**